

網站典藏庫徵集任務之實務與發展現況

Web Archiving Strategies in Collections and Recent Developments

阮紹薇 Shao-Wei Juan

國立臺灣大學圖書館採訪組編審兼採錄股股長

Librarian, Acquisitions Department,

National Taiwan University Library

E-mail: viviany@ntu.edu.tw

【摘要】

本文主要由網站典藏的重要性，介紹世界各國一些重點網站典藏計畫，在實務及學術討論方面的最新發展。其中，網站典藏計畫的策略可分為機構典藏和事件典藏二項重點。其中各計畫針對重大事件的網站典藏，更能彰顯網站典藏的重要性，其徵集內容不但是歷史的一手資料，亦成為典藏單位的館藏特色。超越傳統圖書館的徵集任務與館藏管理，網站典藏庫可提供為文化資產的長久的保存及可即性，提供未來世代研究網站發展、學術議題及歷史進程等重要的研究資源。

【Abstract】

This paper reviews the significance and recent development of web archiving projects and includes a summary of the current state of web archiving in many countries. This paper focuses on NTUWAS as an example of an archive which combines approaches for acquiring information collections, and establishes institutional repository and events-oriented web sites for future studies.

關鍵詞：網站典藏庫；網路資源備份；開放性網路資源；機構典藏；機構網站典藏系統；數位資源寄存法；原生數位資源

Keywords: Web archiving system; Web resource preservation; Open access resources; Institutional repository, IWAS (Institutional Web Archiving System); Digital Collection Deposit Act; Born digital

壹、前言

傳統上，圖書館具有保存人類知識文明的重大使命，在豐富典藏中，呈現著人類發展過程中，所謂「關鍵」、「轉捩點」或「決定性」的重要資訊，記錄著人類歷史及文明發展的重要過程。網際網路興起，資訊科技日新月異，數位資料的產出、傳播和使用也更加便捷，Tim O'Reilly於2004年首次提出Web 2.0的概念，網路內容的提供者或分享平台早已漸漸取代傳統的資訊載體。（Crook, 2009）然而，這些網路空間（website）中，由學術社群間知識分享與學術交流，注重的已不再是內容（content），而是在互動情境（context）激盪的重要意涵，亦有別於傳統圖書館所提供的資訊內容及保存形式。

據World Wide Web（<http://www.worldwidewebsite.com>）的統計，至2010年1月4日，網路中可被搜索引擎索引的網頁數量達到21.24億個所謂表層或淺層的網頁（surface web），然而深藏在網頁背後，必須與網站互動或使用帳號才能檢索到的深度或深層的網頁（deep web），其資訊量更為淺層網頁的500倍之多。在2001年對網路的相關研究，提供一個網頁的平均壽命（lifespan of a web page）約75-100天（Brown, 2006, p.3），在許多文獻中，都引用網站平均生命週期約為44天（項潔、洪筱盈，2006），

更嚴重的是那些足以佐證學術產出的網路引用文獻，被測出可能只有一半的生存機會（half-life）（Seneca, 2009），網站資料稍縱即逝，但影響更深的，就是學術資源的可及性和學術論證的可信度。

近年來，各國政府競相提升國家資訊基礎建設，將文化資產數位化，鼓勵建構學術網路社群及數位產出，越來越多具有學術價值及歷史意義的網路「開放性資源」（open access resources），及只以數位型式產出的「原生數位資料」（born digital）充斥著網路虛擬世界中。如何有效地掌握並有系統的保存、保障這些資源的永久可存取性，以滿足未來可能的研究需求及保存當代的文化與學術資產（洪淑芬，2006），著實考驗著傳統圖書館館藏發展的觀念。

本文將概述網站資源典藏的任務及重要性，並依不同的徵集模式，介紹目前幾個重要的網站典藏計畫。由於大學圖書館肩負著保存知識文化、配合教學研究和推廣學術三大使命，宜建立更完善的網站保存機制，以滿足數位學術產出的管理和保存需求（項潔、洪筱盈，2005），以臺大圖書館於2006年起即開始計劃並執行近四年的「臺灣網站典藏」（NTUWAS）做為經驗分享，本文更實際地探討建構網站典藏計畫的價值及重要性。

貳、網站典藏計畫概述

網站典藏（Web Archiving或Web Preservation）的概念，最早應回溯至網站資料（web）創立之初。Brown（2006）舉World Wide Web Consortium為例，該網站仍保留著1991年的網站型式及內容，將18年前網站的原貌完整保留著（圖1）。若想回顧更多已消失或更新的網站，使用者亦可進入Wayback Machine（網路時光機），為該網站所收1996年10月間的Yahoo!網站（圖2），現在看來與18年前的陽春網站並未有太多變化。

網站流量統計公司Alexa創始人Brewster Kahle於1996年創立了非營利性的Internet Archive（archive.

org，以下簡稱IA），將定期收錄及保存網路中各種文字、音訊、視訊、軟體系統和Web網站資料提供公眾檢閱（圖3）。也是全球第一個由民間成立的網站典藏計畫，使用可透過Wayback Machine鍵入網址，按“take me back”，即可看到IA所保存1996以後網站。IA所典藏的大量網站資料，為學術界留下了網站發展歷史的珍貴記錄，提供了分析網站發展的過程，與網站內容建構的歷史（Brugger, 2009）。

IA的開創性作法，引起了世界各國圖書館、檔案館等資料單位的重視。美國國會圖書館（Library of Congress，以下簡稱LC）在美國政府的強力支援下，即與IA共同開發以「智慧女神」Minerva命名（全



圖1 World Wide Web Consortium
<http://www.w3.org/>



圖2 1996年10月17日Internet Archive第一筆典藏Yahoo!首頁
<http://web.archive.org/web/19961017235908/http://www2.yahoo.com/>



圖 3 Internet Archive 首頁呈現其所收錄的各類型的網站資料

<http://www.archive.org/>

名 Mapping the Internet Electronic Resources Virtual Archive) 的網站典藏計畫，以為將來的世代保存當下的國家文化和智慧產出為己任。

(Arms, 2000) Minerva 的成功開發，主要是獲得「國家數位資訊基礎建設與保存計畫」(National Digital Information Infrastructure Perservation Program, 簡稱 NDIIPP) 撥出 500 萬美金做為啟動經費，由此可知美國政府有鑑於保存網際網路中的「原生數位資料」(born digital) 的急迫性。

NDIIPP 計畫為美國國會圖書館主導，結合聯邦政府及非聯邦政府機構，如美國商務部 (US Department of Commerce)、白宮科技政策辦公室 (OSTP)、美國國家航空暨太空總署 (NASA)、美國國家醫學圖書館 (NLM) 及美國國際圖書館電腦中心 (OCLC Online Computer Library Center) 等單位，以保存數位資源的完整性與統整性為目標，可見國家

級的網路資源的典藏，有賴於相關單位的群策群力、分工合作才足以達成任務。如 NDIIPP 除了 LCWA 計畫外，亦支援加州數位圖書館及紐約大學圖書館等圖書資訊學術單位針對學術網路資源流失進行的 Web-At-Risk Project，建立 Web Archiving Service (WAS) 系統。(Seneca, 2009)

Minerva 於 2008 年改名為 LCWA (Library of Congress Web Archives)，但以所開放公眾瀏覽的網站典藏來看，雖然外型上沒有太多改變 (如圖 4)，然而其典藏工作仍持續進行中。其工作團隊為美國的網站典藏制定相關的技術標準和規範，訂定長期保存的措施，也帶動美國圖書館界重視網站典藏，列入數位館藏徵集的重點開發項目，國際級的圖書館合作組織如 IFLA 及 OCLC 等會員團體，也因此而開始合作研擬相關網站建置及典藏格式 (如 metadata) 等國際標準。



圖 4 美國國會圖書館網站典藏系統LCWA首頁

<http://lcweb2.loc.gov/diglib/lcwa/html/lcwa-home.html>

由LCWA看來，國家圖書館素以典藏國家文化遺產為己任，主導該國的網站典藏計畫當然是無可取代。

（Shiozaki & Eisenschitz, 2009）為了有效收集保存網站資料，自1996年起，如澳洲、奧地利、加拿大、丹麥、芬蘭及日本等國家，相繼開發國家型的網站典藏系統。其中澳洲國家圖書館自1996年起研發網站典藏系統，並將其成功建置的網站典藏系統PANDORA，推向歐盟各國和英國，知名度和規模甚至遠超過美國的LCWA。

2003年7月，歐美12個國家的國家圖書館，共同組成「國際網站典藏合作組織」（International Internet Preservation Consortium, 以下簡稱IIPC），以netpreserve.org為網站名稱，訂定四大目標如下：

1. 收集、保存和徵集全球各機構豐富的網站資源；
2. 開發網站典藏應用程式，研擬國際

共通的網站典藏技術及標準；

3. 成為強而有力的諮詢及遊說團體，向各國政府推廣網站典藏相關立法及政策的制定；
4. 鼓勵及支援各國圖書館、檔案館、博物館及文化遺產等機構網站內容的收集及保存。

IIPC的成果包括開放原始碼（open source）軟體的開發及修正；提出網站典藏標準、詮釋資料（metadata）與內容管理的發展計劃；分享成員組織處理數位內容的經驗；並提供網站典藏的最新動態與未來發展。

自2005年起，IA也開始為個別機構推出年費制的Archive-It（圖5），為110多個合作夥伴提供網站典藏規劃、下載與檢索平台的服務。Archive-It的會員來自於國家圖書館、檔案館、大學圖書館、聯邦機構、非政府非營利機構、博物館，及學者專家等。Archive-It視成員需求，依網址、主題及機構選擇典藏目標，設定徵集範



圖 5 Archive-It網站首頁

<http://www.archive-it.org/>

圍，並運用不同時段的排程下載機制，為會員們定期抓取及保存網站，並提供如matadata的編目資料、全文檢索、網站典藏相關報表及網路內容分析等全面性的服務。

在學術及實務的研討方面，2001年「國際網站典藏工作坊」（International Web Archiving Workshop，以下簡稱IWA）也應運而生，自2001至今，每年召開會議，聚集來自世界各地的相關學者及實務工作者，討論網站典藏的最新發展，分享網站徵集、維護與保存的實務經驗，期能結合學術研究和實務，推動網站典藏工作，以利相關計畫之永續經營。2008年IWA在丹麥召開會議，臺灣方面由時任臺灣大學圖書館系統資訊組主任的陳光華教授派員與會，現場發表學術論文並與來自各國的學術及實務專家們分享「臺灣網站典藏庫」（NTUWAS）的成果，獲得熱烈的回響。

IWA於2009年9月30日及10月1日在希臘召開會議，會議所討論的議題（<http://www.dlib.org/dlib/november09/rauber/11rauber.html>）包含網站典藏技術研發及典藏內容的最新發展，本屆會議多篇論文針對網站典藏源自於機構網站典藏（IWA，Institutional Web Archiving System）概念，討論機構典藏在制定收錄政策、取得授權、及內容完整性等非技術層面的典藏問題。如日本University of Tsukuba即提出機構典藏中必須對網站內容提供者，事先規劃嚴謹的網站典藏政策，以確保公開資料的永久合法性及可用性。

參、網站典藏徵集模式

傳統上，不同類型及規模的圖書館，針對其特殊任務與服務對象，定有不同的館藏發展或館藏選擇政策；在面對如此龐大及來源廣泛的網站資

料，重覆率高、格式不一、內容質量多半未受管控的網路資源，圖書館也必須考量運用有限的資源，制定適合的徵集策略及選擇原則。以下就澳洲的網站典藏計畫PADI（Preserving Access to Digital Information）歸納出的四大典藏徵集模式：全面性、選擇性、專題性、出版寄存式典藏模式，及採取彈性處理方式的綜合性典藏模式，介紹若干國際知名的網站典藏計畫。

一、全面性（Comprehensive Approach）或全域式典藏（Whole Domain）

世界最大，也是唯一企圖全面性（comprehensive approach）典藏全球網站的世界級網站典藏計畫當然首推Internet Archive，使用HTTrack（網站抓取程式web crawler），以自動、批次、全面性的搜尋方式來抓取網站，但因缺乏品質控制、未經篩選即進行儲存，雖有豐富網站資料，但重覆性高，充斥著無計其數失效或不完整的典藏，為其美中不足之處。

此外，國家級的網站典藏計畫以國家或區域性的網站為徵集目標，屬全面網域式典藏模式。瑞典的Kulturarw3和芬蘭的EVA project即為採取這種典藏模式的代表國家。

瑞典皇家圖書館（Royal Library）於1996年開始執行Kulturarw3 Project，修定原有的瑞典

寄存法（Swedish Deposit Law），確保網站典藏的合法性，並利用程式自動抓取屬於瑞典的所有網站資料。可惜的是Kulturarw3在1998年開始建構時，由教育部（The Swedish Ministry of Education）提出立法建議，區隔保存（preserving）和公眾取用（giving access）服務（Mannerheim, 2000），因此至今只限於瑞典皇家圖書館內提供學術研究者申請檢閱使用，Kulturarw3並未向網路大眾開放。

二、選擇性典藏（Selective）

選擇性的網站典藏，意指典藏機構設定特定的網站典藏目標，依主題、重要性及典藏品質等，規劃網站收錄的原則。1996年澳洲國家圖書館建立亞洲第一個網站典藏系統：「潘朵拉」（Preserving and Accessing Networked Documentary Resources of Australia，以下簡稱Pandora），公認為研究並開發國家級網站典藏計畫的先驅，Pandora採用的是選擇性的網站典藏模式，針對澳洲地區的網路出版品和重要網站資源，重點式地典藏具有長期研究價值的網站資源（如圖6）。

Pandora對其所典藏的網站，比照國家圖書館的館藏，利用圖書館分類編目工具，如AACR2、LCSH及DDC等，以半人工方式製作MARC格式，將網站典藏匯入國家圖書館目錄及全國書目中心提供讀者檢索。將



圖 6 澳洲PANDORA系統選擇收錄澳洲區域性網站的各類主題

<http://pandora.nla.gov.au/>

六個州立合作館的網站典藏視為分館的館藏，使用者可從合作館代碼清楚分辨網站典藏的原儲存館，如State Library of Victoria典藏的網站，即在MARC856後加上分段“\$zArchived by ANL for VSL”。（Phillips, 2004）

Pandora建立一個由國家圖書館主導，協調國內重要圖書館，訂定網站典藏的選擇原則（selection guidelines），再將各館所收錄的網站典藏Pandora自行研發的Pandas（Pandora Digital Archiving System）整合成為一個資料庫系統，並結合國家書目系統提供服務，建構一個收錄完整並管理完善的「國家圖書館」網站典藏模式，為各國相關計畫的表率，英國網站典藏庫即應用Pandora系統建置而成。

三、專題式典藏（Thematic）

設定特定主題或事件議題來收

集網站的徵集策略，該計畫組織通常具有特殊的服務對象或任務，並能提供機動性的網站下載功能。美國國會圖書館的LCWA即為專題式典藏的典範，設定特定主題，收藏美國政府及人民所關注的網站資料，如911事件網站（如圖 7）、2002年冬季奧林匹克運動會，伊拉克戰爭及非洲蘇丹的達佛（Darfur）事件等。專題下再依網站主題（subject）、題名（title）、人物或組織名稱（name）提供瀏覽清單。

事件（events）是最能發揮專題式網站典藏計畫的重要性，「911事件」於2001年9月11日發生，為美國歷史上最慘痛的恐怖組織攻擊事件，LCWA自2001年9月11日至2001年12月2日，共收錄2,313個網站（number of sites），同時也將LCWA的合作單位IA中所收錄約3萬個相關網站列入相關資源。此外，2003年非洲蘇丹西部區域發生的種族滅絕事件，世稱



圖 7 美國國家圖書館LCWA的911事件之網站典藏

「達佛危機」（Crisis in Darfur），2005年MINERVA將為「達佛危機」請求國際人道救援的網站或部落格，並抓取（crawl）有關種族滅絕、人道主義干預、維持和平部隊及蘇丹國的相關網站，提供完整的歷史證據及學術研究的資源。（阮紹薇，2009）直至今日，蘇丹達佛地區的人道危機，仍然持續是美國民眾憂心及聲援的焦點。

四、出版寄存式典藏（Deposit）

國家級的網站典藏計畫，多半都必須透過政府立法，明確規定網站所有者將網站自動呈繳至國家的網站典藏庫中。歐洲國家網站典藏計畫最早於1996年開始，即為將網站典藏制定法定寄存法（Legal Deposit Act）或呈繳制度，以有效典藏網路資源，如丹麥、芬蘭、荷蘭、瑞士等。

其中，2008年「國際網站典藏研討會」（IWAU）的召集國家丹麥，

即於2004年12月開始依出版品寄存法（Act on Legal Deposit of Published Material no. 1439 of 22）開始收集所有丹麥國家網域內所出版的國家的或國際的「丹麥資料」（.dk下的所有資料Danish Material），計劃四年為期將全部網站至少典藏一遍，但重要或事件性的重態網站則視需要不定期的選擇性（selective）典藏。（Jacobsen, 2008）

而2009年改版令人耳目一新的英國網站典藏庫（UK Web Archive）（如圖 8），即在2003年制定「英國法定送存圖書館法」（The Legal Deposit Libraries Act，簡稱LDAP）將網站典藏庫比照出版品寄存原則，由British Library, the National Library of Scotland, the National Library of Wales, the Bodleian Library, Oxford, Cambridge University Library and Trinity College Dublin等六所國家及大學圖書館合作網站典藏的寄存任務。至2009年底，總收錄將近1.28億網



圖 8 2009年最新改版的英國網站典藏庫（UK Web Archive）

<http://www.webarchive.org.uk/ukwa/>

頁，企圖完整收藏英國的所有網站。由於收錄網站內容可能會衝擊到文化事業、運動或媒體等有經濟價值的版權資料，因此部份內容只限六家寄存圖書館的閱覽室開放檢閱。（Tuck, 2009）

美洲地區的加拿大於2004年開始進行Library and Archives（以下簡稱LAC）的網站典藏計畫，由加拿大圖家圖書館和國家檔案館依「加拿大圖書館與檔案法」（The Library and Archives of Canada Act, the Legislative Act）的規定，收錄加拿大境內政府（*.gov）及私人網站。該計畫從為網站典藏的特性，研擬編目和描述網站存取層次（access level）的原則，並設定數位典藏發展政策（digital collection development policy）、網站選擇與徵集指導原則（selection

and acquisition guidelines for Canadian web sites），及數位出版品的記述原則（description policy for digital publications），將網站典藏視為傳統的紙本出版品，建立書目記錄並進行後續典藏徵集作業。（Lilleniit, 2007）由於寄存制度的法源及支持，LAC執行二大計畫：E-Collection和Government of Canada Web Archive，將政府網站及網路出版資源完整保存下來，為所有典藏建立專業編目紀錄，並開放公眾檢索取用。

五、綜合性典藏策略（Combined Approaches）

由於各種網站典藏模式均有利有弊，部份國家採取綜合性的策略，選定特殊事件或議題做全域式的收集，

對其他的網站採取較為嚴謹的選擇性典藏原則。如前述丹麥和英國網站典藏計畫（UK Web Archive）的作法。在亞洲方面，2001年起中國大陸、日本及韓國亦採取這種綜合性典藏策略進行網站典藏。

中國國家圖書館於2001年成立「網絡信息資源采集與保存實驗項目」（Web Information Collection and Preservation，以下簡稱WICP），初期由北京大學開發「中國Web信息博物館」（以下稱Web InfoMall）的網站典藏系統。WICP計畫於2003年開始啟動，以長期追蹤並保存政府官方網站及電子報為首要典藏目標，再配合「專題存檔」為典藏策略，如「奧運專題」全面收錄2008年北京奧運共60個子專題的相關網站，完整記錄運動場地建設、運動項目，和奧運吉祥物等。但由於中國國家圖書館未對外開放其網站典藏的成果，目前也只能從Web InfoMall輸入正確的網址檢索其網站典藏的情況。

同樣身處於亞洲地區的國家級網站典藏計畫，還有日本國會圖書館執行的WARP（Web ARchiving Project）計畫及韓國國家圖書館的OASIS（Online Archiving & Searching Internet Sources）。韓國的網站典藏計畫則較近似於將數位典藏的網站透過網站典藏計畫加以保存，重視國內的數位出版內容的網站典藏。日本國會圖書館的WARP主要特色為遵照「公眾送信權」（Rights of Public

Transmission，引自日本國著作權法23條），所有典藏均有詳細的審核和授權機制。WARP收錄中央及地方政府網站資料外，亦保存網路上免費公開的電子期刊與圖書，此外，WARP也接受推薦，避免遺漏重要典藏。如圖9，WARP將日本新潟越後妻地區都會每年舉辦的『大地藝術祭』相關網站保存下來，提供背景分析及描述網站的「書誌情報」。

在臺灣方面，洪淑芬（2006）發表國內第一篇完整介紹網站典藏概念的論文，即主張基於完整性與經濟效率，宜考慮混合式之模式，除了批次選擇收集特定類別網站（如.edu、.gov、.org等），亦個別選擇已知的重要網站作深入的蒐集。而國家圖書館身為國家級圖書館，除致力實體文獻的數位化及資訊服務，也注意到網路資源逐步取代實體文獻之趨勢，在進行國家文獻典藏時，網路資源的典藏也是不能忽略的重要職責，基於資源保存與學術研究的立場，國家圖書館於2007年著手建置「國家圖書館網站典藏先導系統」，目前已開放公眾檢索使用。該網站提供多元化的檢索查詢方式，如典藏查詢、主題瀏覽、機構瀏覽、地區瀏覽及重大議題瀏覽，亦歡迎讀者上網推薦。

以下筆者就臺灣大學圖書館建置之「臺灣網站典藏庫」（NTUWAS），簡述該系統所收錄之館藏特色，就實務分享網站典藏的重要意義及價值所在。



圖 9 WARP中日本『大地藝術祭』相關網站的典藏

肆、臺灣大學網站典藏庫的館藏特色

2006年臺灣大學圖書館承教育部委託，建置「臺灣機構典藏」以存取並長久保存機構知識的產出，提出加強收集與管理機構內產生的各種數位資源及灰色文獻，提高機構學術成果的能見度與影響力。（項潔、洪筱盈，2005）2006年7月，臺大圖書館結合系統資訊組及「臺灣機構典藏」計畫的工作團隊，將原生數位型態（born digital）的「網站典藏」（Web archives）列入「臺灣機構典藏」的計畫項目，將長期關注相關計畫的研發成果，成功建置臺灣第一個網站典藏系統NTU Web Archive

System（以下簡稱為NTUWAS）（如圖10）。

NTUWAS的系統管理介面主要有「後端管理系統」及「前端讀者介面」二部份。後端設定網站下載及存取機制，不同時期的網站典藏能有效對應讀者查詢介面，管理者亦能清楚的分辨尋找網站典藏，增進管理的效能。而「前端讀者介面」除提供網站檢索及全文檢索，讀者可提供本系統未收錄之網站逕行推薦，讓網站內容更加完備。NTUWAS系統建置及架構的詳細介紹，可參考陳光華教授領導的團體於2008年發表於第八屆“International Web Archiving Workshop”的論文“Developing National Taiwan University Web Archiving System”。（陳光華、陳彥



圖 10 2006年12月建置初期的NTUWAS

良、丁培峰，2008）

NTUWAS的內容建立，如同實體館藏之徵集，由圖書館採訪組規劃館藏發展方向，並實際執行網站之選擇與下載。經考量如Internet Archive系統自動下載造成大量無效網站的缺點，NTUWAS主要採用人工篩選，再輔以系統排程功能的收錄原則，最後再經由網站檢測確保收錄品質，才提供讀者瀏覽與檢索。

NTUWAS前後台系統功能經過二年的測試，於2008年4月正式上線供眾使用，同時利用圖書館快訊並登刊於臺大校訓，歡迎全校師生共同投入這項重要的網站典藏工作。目前NTUWAS已收錄近5,000種網站資料，使用12T的硬碟空間，提供更多元化及有趣的各樣瀏覽典藏的介面，如觀察網站更迭的「時光迴廊」、依

事件發生時間瀏覽典藏的『時光拉霸機』和將地域性網站以Google地圖呈現的『地理瀏覽』等。（如圖11）

NTUWAS為臺灣第一個網站典藏系統，鎖定選擇、蒐集、保存原始性網頁資料目標，善盡保存臺灣網站數位資產之責任，並提供綜存網站新舊版本的展示平台及檢索機制，以滿足當代與未來世代研究發展之需求。然而，網站資料豐富且瞬息萬變的網路世界中，NTUWAS如同全世界各國的網站典藏計畫一般，都不可能做到一網打盡式的目標。因此在NTUWAS的典藏方面，工作小組即參考各國網站典藏系統的作法，設定計畫目標如下：

1. 提供網站典藏（Web archive）機制；
2. 記錄網站發展的歷程；



圖 11 至2009年多次更新改版後的NTUWAS
<http://webarchive.lib.ntu.edu.tw/default.asp>

3. 收藏網路記憶中的臺灣紀錄，提供未來學術研究的參考。

然而，網路資料超越時間及空間的限制的特性，在網站的典藏方面也必須考量全球性學術、政治、經濟貿易上互相依存、不斷增強的脈動關係，因此除了保存臺灣這片土地上人民和國家的網站典藏外，也應對應、觀照同時期國際社會的變化脈絡，NTUWAS也運用有限的資源，盡可能將臺灣所關注的重大國際事件及學術網站等列入典藏目標。

綜合以上，NTUWAS網站典藏以臺灣大學圖書館機構典藏目標為優先，並規劃富有地方性及時代價值的特色網站達到保存臺灣記錄的計畫目標。藉此設定了NTUWAS三大典藏發展目標：典藏機構、典藏臺灣和典藏時代，並以內而外的館藏發展策略，

逐步規劃收錄內容及典藏範圍。

以下就NTUWAS的三大典藏發展目標，以實際的網站典藏舉例說明如下：

一、典藏機構

NTUWAS的計畫主目標在於將臺灣大學所有及相關機構的網站資料，完整保存；並以臺大身為教育部所屬國立大學的機構性質，擴大收錄公部門網站及教育學術相關網站，期能至少完整收錄一遍，並儘可能典藏重要的網站更迭紀錄。

（一）臺大相關網站

完整收藏臺灣大學首頁、行政及學術單位、合作網站、計畫網站，亦收錄重要的個人網站。典藏特色如

下：

1. 臺灣大學首頁的網站典藏回溯至1997年7月15日，至2010年1月6日止，典藏數量為147個，並持續記錄臺灣大學所屬單位的更新網站，或重大異動事件。如圖12，NTUWAS保存了「臺灣大學進修推廣部」學分班停招前的網站。
2. 收錄臺灣大學2008年11月17日80週年校慶『八十臺大、前進百大』的相關網站，典藏「邁向頂尖大學計畫」補助的新增計畫或單位之相關網站。並擴大收錄其他獲得補助的大學相關計畫網站，及接受繁星計畫的重點高中網站，作為未來教育政策研究的參考。

（二）政府機構網站

中央各部會、各級地方政府所屬單位之合作網站及個別網站。原則上中央及各部會網站因為過於龐大，因

此主要收錄總統府、五院首頁及行政院下重要二級單位網站，重要的館藏特色如下：

1. 2007年初臺灣興起一股臺灣正名風潮，部份政府事業單位更名而進行的網站改版：如中華郵政股份有限公司於2007/02/12更名為臺灣郵政股份有限公司，又於2008/08/01回復名稱為中華郵政股份有限公司（圖13），又如中正紀念堂管理2007/05/15更名為國立臺灣民主紀念館，又於2008/08/21回復名稱為國立中正紀念堂管理處，一波三折的過程；亦有更名後維持新名稱的，有中國石油股份有限公司改為臺灣中油股份有限公司，中國造船改名為臺灣造船等。NTUWAS典藏並保存與這些因單位名稱改變而更迭的網站，也無形中成為見證臺灣歷史的一手資料。
2. 臺灣在2008年進行的立法委員與總



圖 12 臺灣大學進修推廣部學分班停招前的網站

http://webarchive.lib.ntu.edu.tw/switch.asp?IDX_SITE=1197&LOADDTE=2008-05-01



圖 13 2007年2月12日臺灣郵政股份有限公司的網站

http://webarchive.lib.ntu.edu.tw/switch.asp?IDX_SITE=1046&LOADDTE=2007-02-13

統大選，從候選人網站或部落格的網站典藏，可說是在網路時代出現新形態的選舉策略。由NTUWAS典藏的競選相關網站來看，約有三分之一個網站已大幅變更或不再運作，因此必須掌握時機保存新聞性或時效性之網站。中央及地方選舉後的各級政府人事或網站異動的記錄。重要如『阿扁電子報』全部內容（圖14）及其活動紀錄（圖15）。

又如2008年5月21日總統大選前後總統府網站的異動，如圖16為馬英九總統上任後新架的總統府網站，經由NTUWAS所設計由人工操作的「即時下載」功能，才能留下如此珍貴的網站歷史記錄。

3. 中華民國臺灣省319鄉鎮及55臺灣原住民鄉鎮的地方網站，保留豐富的地方文史資料及地方產業特色等，這部份亦可視為達成第二個典藏目標「典藏臺灣」的重要典藏特色。如高雄縣三民鄉於2008年1月1日改名為那瑪夏鄉，及2009年8月8

日的八八水災，及2010年3月4日甲仙百年大地震受災地區，留住重要關鍵歷史紀錄的地方政府網站。

4. 臺北市政府社會局廣慈博愛院於民國96年9月11日正式被裁撤，但舊網站仍存在NTUWAS中。

（三）教育學術網站

主要收集重點國立或私立大專院校及學術單位之網站、合作網站（結盟網站）、計畫網站及個人網站等。

典藏特色主要為：

1. 前述「邁向頂尖大學計畫」相關學校網站；
2. NTUWAS也致力肩負學術機關的任務，配合臺灣大學正在進行的邁向頂尖大學計畫為基礎，兼收藏校外其他相關主題網站，典藏相關的學術成果，關於這方面，其特色可分為：地區性研究主題如亞太研究、南島文化，及臺灣即將邁入高齡化社會衍生之公共保險研究等；將文化資產加值利用的數位典藏計畫；高科技的奈米科技、光電與半導體



圖 14 中華民國總統府網站『阿扁電子報』

http://webarchive.lib.ntu.edu.tw/switch.asp?IDX_SITE=96&LOADDTE=2006-08-11



圖 15 中華民國總統府網站陳水扁總統的活動紀錄

http://webarchive.lib.ntu.edu.tw/switch.asp?IDX_SITE=96&LOADDTE=2008-05-20

等；探索人類生命奧秘的基因體與分子生物等；及全球性環境保育及生物多樣性等主題。

3. 重要異動機構：原國立體育學院於2008/02/01更名為國立臺灣體育大學（桃園），並於2009/08/01更名為國立體育大學的網站紀錄。在NTUWAS「時光迴廊」中點選右方的「教育學術」類目，即可一探該

校更名前後的網站變化（圖17）。

二、典藏臺灣

（一）相關網站

文學、美術、古物、民間風俗、民間藝術等及相關典藏單位（如博物館、美術館）網站。各教學資源及史



圖 16 中華民國總統府網站首頁馬英九就任第12任總統的相關活動

http://webarchive.lib.ntu.edu.tw/switch.asp?IDX_SITE=96&LOADDTE=2008-05-21

料檔案等相關網站入此。NTUWAS保存臺灣重要藝術文化網站共345種，其中藝術與文化資產部份265種網站及華文及臺灣文學網站98種，NTUWAS保存臺灣藝術文化資源的重要典藏特色。

（二）族群相關網站

原住民、客家、臺灣各宗親會網站；婦女團體、少數民族、弱勢族群入此。這部份的重要典藏如：

1. 保存臺灣原住民族相關網站及部落格，保存2007年來許多消失或停止運作中的臺灣原住民族網站，如蕃薯藤網站中的「蘭嶼曙光」（圖18）。
2. 華僑及臺灣新移民等相關網站，如圖19「外籍及大陸配偶家庭關懷服

務網」，為階段性任務並已結束的政府網站。

（三）非政府組織

職業公會及工會、慈善團體、宗教團體（含臺灣傳統民間信仰及民俗團體）、其他非營利社團等。這部份呈現臺灣社會的良善及公益的一面，也是最能反映臺灣大學的社會關懷。由於非政府組織的種類繁多，NTUWAS針對自然資源保護、婦女與性別相關及人道救援三項，與學術研究或時事議題較相關的重要議題組織，設立獨立類目，以觀察相關議題及網路資源之發展。如圖20「臺灣國際醫衛行動團隊」（Taiwan International Health Action，簡稱TaiwanIHA）



圖 17 NTUWAS「時光迴廊」中，國立體育學院二度更名的網站紀錄

<http://webarchive.lib.ntu.edu.tw/machine.asp?selvalue=008001000>



圖 18 已消失的「蘭嶼曙光」

http://webarchive.lib.ntu.edu.tw/switch.asp?IDX_SITE=1916&LOADDTE=2007-03-05



圖 19 已消失的「外籍及大陸配偶家庭關懷服務網」

http://webarchive.lib.ntu.edu.tw/switch.asp?IDX_SITE=820&LOADDTE=2007-01-10

三、典藏時代

如前言所述，吾可以從圖書館豐富典藏中，找到人類發展過程中，所謂「關鍵」、「轉捩點」或「決定

性」的重要資訊，在這些重要的歷史資訊中，部份重大的事件和重要人物，若以獨立或特別的類目來呈現，一則以呈現其時代意義，一則不會埋沒在龐大的分類目錄中，被人遺忘。這部份可能會因為事件或人物本身後續衍生更多的討論或產生更多網路資源，因此可隨時將後分類進行彈性調整。

(一) 事件網站

做為美國國會幕僚機關的美國國會圖書館，典藏了甚多重要的事件網站，記錄重大歷史事件發展過程，如收錄聯合國和美國國會所關切的『達佛危機』，及2001年「911事件」等相關網站；澳洲國家圖書館的Pandora收錄2004年發生的「南亞海嘯」（Asia Tsunami Disaster）相關區域或國家的網站；英國的UK Web Archiving典藏「倫敦恐怖攻擊」



圖 20 TaiwanIHA網站首頁

http://webarchive.lib.ntu.edu.tw/switch.asp?IDX_SITE=4525&LOADDTE=2008-02-23

(Terrorist Attack-London, 7th July 2005) 等。

從各國網站典藏計畫中的均將事件 (events) 仍為重要典藏類目，並以不受時空及一般類目的網站典藏的處理辦法，所以「事件導向」(events oriented) 的網站典藏，更能讓網站典藏顯出它的重大意義。

中國國家圖書館WICP的Web InfoMall，收錄熱門事件如「抗擊非典的日子」和「2008北京奧運」等相關新聞及活動網站，並「歷史的回放」來呈現這些具有時代意義的事件；日本國會圖書館將不復存在的行政區域舊網站保存在WARP中，則提供了地方性的文史檔案紀錄。(阮紹薇，2009)

NTUWAS為重要專門討論歷史事件或時事議題的網站，如228事件、傳染病 (SARS)、軍購案、921大地震等，列入「重大事件」中，再依

歷史事件、改革議題、爭議問題、全球與環境問題略作分類。而就事件的發生有的在人們的心目中，主要以時間為事件的記憶焦點，如2000年的921大地震、2008年美國總統大選及2009年的88水災等，NTUWAS特別設計「時光拉霸機」，以時間點瀏覽模式，集中呈現事件當下的網站內容 (圖21)。

「時光拉霸機」以時間概念讓使用者可以依時間的記憶，就該事件在NTUWAS中典藏的相關網站查詢當時的網站資訊，亦為NTUWAS的一項創舉，如查詢莫拉克颱風 (八八水災) 當時地方政府利用網站傳遞的受災情形及政府公告 (圖22)、總統府網站的總統勘災現場及時消息，及後續發展的重要紀錄：如監察院於2009年8月21日發布88水災召開專案調查研議小組會議的監察院網站公告等。

除了時間的記憶外，許多議題或事件，因其區域性或別具意義的地理位置，若以地圖概念來呈現，將更有助提示使用者瀏覽網站典藏，因此NTUWAS利用google地圖的免費資源，提供依「地理瀏覽」網站典藏的服務，如以世界地圖呈現的臺灣邦交國的位置，呈現海島邦交國密集的畫面，而由臺灣地圖呈現臺灣原住民鄉鎮 (圖23)，亦有助於臺灣原住民分佈區域的辨識。

由「地理瀏覽」展示網站典藏中地區性的歷史軌跡，亦是NTUWAS的一項創舉。如比較NTUWAS中921地

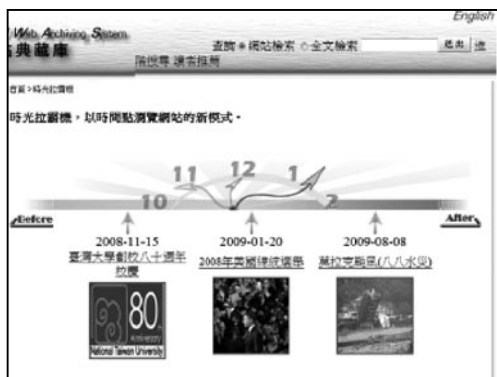


圖 21 NTUWAS的時光拉霸機

<http://webarchive.lib.ntu.edu.tw/history.asp>

震（圖24）及88水災的災區地標（圖25），會發現很大的重疊性，想見1999年921震災過後，2008年10月30日「財團法人921震災重建基金會」因任務告一段落才走入歷史，2009年8月8日部份受災區又遭受風災的侵襲，而這些地區又幾乎都落在我們原住民的鄉鎮。

（二）人物網站

人物網站是在事件網站之外，另一種書寫歷史的重要一手資料。NTUWAS中在各類型網站中都收錄一些人物資料，一些已過世的臺灣重要人物，如李國鼎先生的紀念網站、陳定南先生生前的部落格，各行各業如文學家、藝術家和教師的個人網站，保存重要的傳記資料和學術研究資源。

在NTUWAS保留的人物網站中，



圖 22 NTUWAS時光拉霸機中八八水災專題呈現災區地方政府網站的重要資訊

<http://webarchive.lib.ntu.edu.tw/tevent.asp?eventValue=33>

最熱鬧的要屬臺灣社會中公職人員選舉活動頻繁，而大量產生的競選人物網站。各種候選人的部落格、個人網站和政黨後援的個人網站，尤其是選舉過後，幾家歡樂幾家愁，許多落選人從此就關閉網站，資料永遠消失，則是這部份人物網站的最大特色。

然而這些網站保留著許多政治人物因為選舉宣傳而留下平常難得一見的個人資料，及因為爭取選票而提出的政見，甚至是利多大放送等，這些快速被建立、在活動期間不斷更新的網站（如最新選戰消息），短期內又很可能快速消失的人物網站，也是許多國家在網站典藏中不可缺少的一個徵集重點。

這部份因為涉及Web 2.0的網路

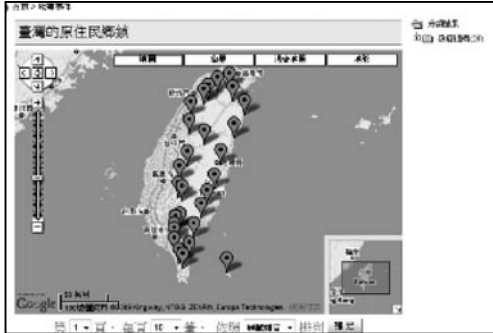


圖 23 NTUWAS依臺灣原住民鄉鎮地理位置的網站典藏瀏覽

<http://webarchive.lib.ntu.edu.tw/gevent.asp?eventValue=1>

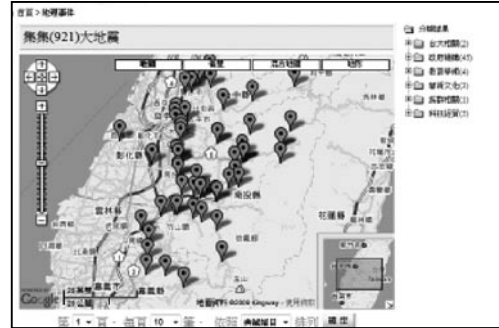


圖 24 NTUWAS依921大地震地理位置的網站典藏瀏覽

<http://webarchive.lib.ntu.edu.tw/gevent.asp?eventValue=8>

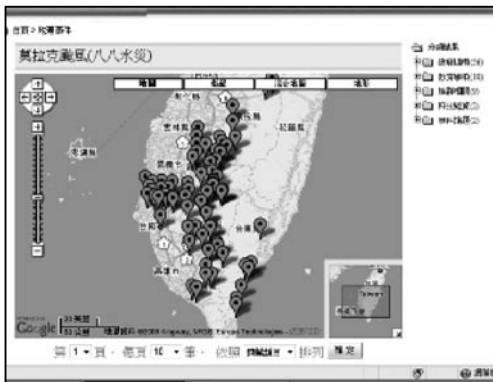


圖 25 NTUWAS依88水災地理位置的網站典藏瀏覽

<http://webarchive.lib.ntu.edu.tw/gevent.asp?eventValue=7>

應用和社群經營模式，目前網站保存工具如HTTRACT無法成功抓取社群網站如部落格及影音分享網站等，也成為許多網站典藏計畫的一大挑戰。（Crook, 2009）

2008年韓國、美國及中華民國都舉行總統大選，NTUWAS均保留當時相關人物及政黨的競選網站。除現任美國總統歐巴馬的個人競選網站外（圖26），也保留了當時和歐巴馬同時角逐民主黨的美國總統候選人的希拉蕊個人網站，及美國民主黨網站；而共和黨代表角逐總統大位的馬侃（John McCain）因其夫人抄襲網站食譜而釀成負面新聞話題的網路事件，也正好被NTUWAS保留下來（圖27）。

2008年中華民國亦舉行影響政黨輪替的總統大選，選情激烈，而各候選人的個人及政黨的競選網站，也相



圖 26 NTUWAS中2008年美國總統候選人Obama最初的競選網站

http://webarchive.lib.ntu.edu.tw/switch.asp?IDX_SITE=4322&LOADDTE=2008-01-24

當精采。當時民進黨總統候選人謝長廷在網站中提供成長歷程中的個人資料呈現感性訴求（圖28）。而在後來順利當選馬英九總統，其www.ma19.net的部落格由來已久，NTUWAS自2006年9月19日開始保存MA19網站，其中2006年11月14日的典藏也以懷念親人為主題（圖29），由此可見政治人物個人官網為提供重要傳記資料和政治經歷的重要來源。

在2008總統大選激情過後，NTUWAS被推薦為馬蕭總統大選的網站存檔處，為NTUWAS導入大量的網站點選數量，讓許多使用者認識臺灣網站典藏庫，在NTUWAS的線上推薦網頁中提供許多值得典藏的網站。（圖31）

伍、結語

在NTUWAS的「時光迴廊」中，回顧不同時間網站的重要改變，如「更名」、「改版」及「關站」等，從臺灣網站典藏庫的館藏來看，網站的變遷反映了臺灣社會和政治的變化，這些變化均被NTUWAS忠實地保存下來。當國立中正紀念堂管理處於民國96年5月15日更名為國立臺灣民主紀念館（如圖30，在改名初期「大中至正」與臺灣民主紀念館並存的首頁）。

保存當下的重要事件，以免事後遺憾，一直都是網站典藏工作的一大挑戰。NTUWAS能掌握到這些重要的網站更迭時機，是靠著「即時下載」後台系統功能；此外，為應付政府機構更名、更換首長或組織重整時，有大幅度的變動，由於這部份的異動網站量大，無法用人工一一控制，因此，NTUWAS也為此特別設計「預約排程」應付大量網站同時更新的下載排程。此為NTUWAS從2006年起不斷在進行平台測試和內容規劃上不斷改進的成果。

當然，由人工進行網站選擇及下載網站品質測試及篩選，NTUWAS會過濾掉那些明顯為娛樂、入口、商業性質，或內容不宜的網站。而因為網站本身使用Flash程式寫作的網頁、Javascript動態網頁連結資料、深層網站（deep web）時，或在網站下載後的測試階段出現dead link、亂碼、file not found（錯誤代碼404）等問



圖 27 NTUWAS 中 2008 年美國總統候選人 McCain 競選網站中被控抄襲的家庭食譜

http://webarchive.lib.ntu.edu.tw/switch.asp?IDX_SITE=4334&LOADDTE=2008-01-22



圖 28 NTUWAS 中民進黨總統候選人謝長廷競選網站呈現感性訴求

http://webarchive.lib.ntu.edu.tw/switch.asp?IDX_SITE=177&LOADDTE=2007-10-26



圖 29 NTUWAS 自 2006 年起保留的馬英九官方網站

http://webarchive.lib.ntu.edu.tw/switch.asp?IDX_SITE=176&LOADDTE=2006-11-14



圖 30 臺灣民主紀念館「大中至正」共存的網站

http://webarchive.lib.ntu.edu.tw/switch.asp?idx_site=3544&loaddte=2007-05-24



圖 31 馬蕭總統大選的網站存檔聲明

http://webarchive.lib.ntu.edu.tw/switch.asp?IDX_SITE=176&LOADDTE=2008-04-03

題，部份重要網站仍會繼續測試直到下載到內容較為完整的網站。

至於網站典藏的類目方面，由於NTUWAS採取依典藏來進行「後分類」，配合收錄網站的性質與數量，以及日後典藏的重點項目，來進行適度的更新或調整。因此除了主要大類不變外，子類目隨時會作變動，也是網站典藏的一項較圖書資料編目較為彈性的作法。

NTUWAS自2008年4月初開放典藏和系統的成果至今，期盼全校師生和各界使用者在典藏內容及主題範圍方面提供意見，讓NTUWAS持續收錄更多值得珍藏、具學術性或有時代價值的網站。使用者除可利用「讀者推薦」功能協助我們新增典藏內容外，亦可點選「聯絡我們」利用電子郵件

來信批評指教。期待NTUWAS未來能比照臺大機構典藏計畫，結合相關教育學術團體，擴充網站資料收錄範圍，規劃蒐集臺灣地區以外各類網路資源，達成提供世界級學術資源的終極目錄。讓臺灣的網站典藏計畫邁向世界級Web Archiving計畫的目標。

本計畫系統建置及維護有賴「臺灣大學機構典藏」主持人項潔教授、陳雪華館長，及圖書館系統資訊組陳光華教授的指導，帶領系統資訊組丁培峰、陳彥良，及陳建智等同仁，先後參與NTUWAS前后台的系統規劃及程式設計，開發期間不斷發揮程式專才及創意，設計出別具特色的檢索平台和瀏覽介面，倍受國際相關計畫組織的重視，被國際網站典藏工作坊（IWA）推薦為重要網站典藏計

畫。在內容建置方面，筆者自2006年接受執行網站典藏徵集的任務以來，除原機構典藏小組全然授權外，採訪

組梁君卿主任數年來大力支持，廖偉辰先生熱心負責，協助內容建置功不可沒，特此致上最深的謝意。

參考網站

日本國會圖書館「国立国会図書館インターネット情報選択の蓄積事業」（簡稱WARP）。上網日期：2010年1月10日，網址：<http://warp.ndl.go.jp/>

中國北京大學「中國Web信息博物館」（簡稱Web Infomall）。上網日期：2010年1月10日，網址：<http://www.infomall.cn/>

英國國家圖書館“UK Web archiving consortium”（簡稱UKWAC）。上網日期：2010年1月10日，網址：<http://www.webarchive.org.uk/ukwa/>

美國國會圖書館“Library of Congress Web Archives”（簡稱LCWA）。上網日期：2010年1月10日，網址：<http://lcweb2.loc.gov/diglib/lcwa/html/lcwa-home.html>
<<http://lcweb2.loc.gov/diglib/lcwa/html/lcwa-home.html>>

國家圖書館「臺灣網站典藏先導系統」。上網日期：2010年1月10日，網址：<http://webarchive.ncl.edu.tw/nclwaFrontNew/>
<<http://webarchive.ncl.edu.tw/nclwaFrontNew/>>

臺灣大學「臺灣網站典藏庫」（簡稱NTUWAS）。上網日期：2010年1月10日，網址：<http://webarchive.lib.ntu.edu.tw/>
<<http://webarchive.lib.ntu.edu.tw/>>

澳洲國家圖書館“Pandora Digital Archiving System”（簡稱Pandora）。上網日期：2010年1月10日，網址：<http://pandora.nla.gov.au/>

IFLA/IPA Steering Group. Preserving the Memory of the World in Perpetuity: a joint statement on the archiving and preserving of digital information。上網日期：2010年1月10日，網址：<http://archive.ifla.org/V/press/ifla-ipa02.htm>

更多網站典藏計畫相關網站請見圖 30二。上網日期：2010年1月10日，網址：<http://webarchive.lib.ntu.edu.tw/webs.asp>
<<http://webarchive.lib.ntu.edu.tw/webs.asp>>



圖 32 NTUWAS 中各國網站典藏系統的簡介及連結

<http://webarchive.lib.ntu.edu.tw/webs.asp>

參考文獻

- 洪淑芬（2006）。網路資源備份保存技術與應用探析。*國家圖書館館刊*，95（2），75-121。
- 項潔、洪筱盈（2005）。臺灣機構典藏發展芻議。*教育資料與圖書館學*，43（2），173-190。
- 項潔、洪筱盈（2006）。大學圖書館數位保存與館藏發展策略：以臺大圖書館為例。*大學圖書館*，10（2），2-18。
- 陳光華、丁培峰（2007）。臺灣大學網站典藏系統之研發。*大學圖書館*，11（2），1-18。
- 阮紹薇（2008）。從網站典藏臺灣—「臺灣網站典藏庫」。*臺灣大學圖書館館訊*，新96期。
- 阮紹薇（2009）。回溯時光 典藏記憶—談網站典藏庫的事件記錄。*臺灣大學圖書館館訊*，新113期。

- Arms, W. Y. (2000). *Collecting and Preserving Open-Access Materials on the Web*. Retrieved March 11, 2008, from <http://www.cs.cornell.edu/wya/LC-web/proposal.html>.
- Brown, A. (2006). *Archiving websites: a practical guide for information management professionals*. London: Facet Pub.
- Brugger, N. (2009). Website history and the website as an object of study. *New Media & Society*, 11(1&2), 115-132.
- Cantello, G. & Stegenga, J. (2008). Government web content in Canada: a national library web archive perspective. In *World Library and Information Congress: 74th IFLA General Conference and Council*. Retrieved May 27, 2008, from <http://www.ifla.org/iv/ifla74/index.htm>.
- Chen, K. H., Chen, Y. L. & Ting, P. F. (2008). Developing National Taiwan University Web Archiving System. In *Proceedings of the 8th International Web Archiving Workshop*, Aarhus, Denmark: University of Aarhus.
- Crook, Edgar (2009). Web archiving in a Web 2.0 world. *Electronic Library*, 27(5), 831-836.
- Jacobsen, G. (2008). Web Archiving: Issues and Problems in Collection Building and Access. *Liber Quarterly: The Journal of European Research Libraries*, 18(3/4), 366-376.
- Linneniit, R. (2007). Archiving the Canadian web: experiences at Library and Archives Canada. *The Serials Librarian*, 53(1/2), 139-149.
- Mannerheim, J. (2000), *Preserving the Digital Heritage of the World*. Retrieved December 29, 2009, from <http://etjanst.hb.se/bhs/ith/1-00/jm2.htm>
- Masanès, J., & Rauber, A. IWAW' 06 proceeding [sic] of the 6th International Web Archiving Workshop, September 21-22, 2006, Alicante, Spain Available from <http://www.iwaw.net/06/PDF/iwaw06-proceedings.pdf>.
- Paynter, G., Joe, S., Lala, V., & Lee, G. (2008). A Year of Selective Web Archiving with the Web Curator at the National Library of New Zealand. *D-Lib Magazine*, 14(5/6), 2-2.
- Phillips, M. E., & Koerbin, P. (2004). PANDORA, Australia's Web Archive: How Much Metadata Is Enough? *Journal of Internet Cataloging*, 7(2), 19-33.
- Rauber, A., & JulienMasanes. (2008). Report on the 8th International Workshop on Web Archiving - IWAW 2008. [Proceeding]. *D-Lib Magazine*, 14(11), 14-14.

- Seneca, T (2009). The web-at-risk at three: overview of an NDIIPP Web Archiving Initiative. *Library Trends*, 57(3), 427-441.
- Shiozaki, R. Y. O., & Eisenschitz, T. (2009). Role and Justification of Web Archiving by National Libraries: A Questionnaire Survey. *Journal of Librarianship & Information Science*, 41(2), 90-107.
- Tuck, J. (2008). Web Archiving in the UK: Cooperation, Legislation and Regulation. *Liber Quarterly: The Journal of European Research Libraries*, 18(3/4), 357-365.